



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

Address: COMMISSIONER FOR PATENTS

P.O. Box 1450

Alexandria, Virginia 22313-1450

www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/534,661	11/10/2005	Liang He	42P14283	6515
45209	7590	04/21/2011	EXAMINER	
MISSION/BSTZ			GUERRA-ERAZO, EDGAR X	
BLAKELY SOKOLOFF TAYLOR & ZAFMAN LLP			ART UNIT	
1279 OAKMEAD PARKWAY			PAPER NUMBER	
SUNNYVALE, CA 94085-4040			2626	
MAIL DATE		DELIVERY MODE		
04/21/2011		PAPER		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary

Application No.

10/534,661

Applicant(s)

HE, LIANG

Examiner

EDGAR GUERRA-ERAZO

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 02/08/2011.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-49 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1, 4, 6, 8, 10, 14, 35, 38, 45-49 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-040)
- 3) ☐ Information Disclosure Statement(s) (PTO/SB-08)
Paper No(s)/Mail Date _____
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: _____

DETAILED ACTION

Response to Amendment

1. In response to the office action from 12/08/2010, the Applicant has submitted an amendment filed on 02/08/2011, amending Claims 1, 35, 38, and 45, and arguing to traverse the art rejection based on the remarks at Pages 9-11 of the Amendment.
2. Applicant's arguments have been fully considered but they are not persuasive. The previous rejection is maintained, altered with respect to the amended claims and due to the reasons listed below in the response to arguments.
3. In response to the amendment of Claim 35, the previous rejection under 35 U.S.C. 101 has been withdrawn as Claim 35 is now limited to a non-transitory computer-readable medium, and thus transitory or signal-based mediums are excluded.

Response to Arguments

4. Applicant's arguments have been fully considered but they are not persuasive for the following reasons:
5. With respect to Claims 1, 35 and 45, the Applicant appears to argue at Page 11 of the Amendment that none of the cited references, individually or when combined, each reasonably suggest retrieving speech recognition information from the internet as recited by the amendment in claim 1, and furthermore that because of these reasons the cited references also do not teach or reasonably suggest "performing speech processing based on the selected web interaction mode and the retrieved speech processing information". In response, the Examiner respectfully notes that for instance at Col. 11, Lines 17-60 of the Polcyn reference, it is disclosed

how a message may comprise various different types of data such as a segment may comprise different types of data communicated to EMS 206 where the message may be thought of as a “container” in which various media are generated and organized into segments of different types. Furthermore, it is also disclosed in Polcyn how web interaction modes can be used independently or concurrently in an order of communication where for instance a first segment may comprise information communicated from a communicating party in a first communication session to EMS 206, while a second segment may comprise information communicated in a second session to EMS 206 in both instances the messages comprising various different types of data. At Col. 12, Lines 53-65 and Col. 10, Lines 6-36, of Polcyn for example, it is also disclosed how the plurality of web interaction modes encountered in the plurality of segments of the plurality of messages in the communication interactions are able to work in agreement with a transcription interface application that may automatically transcribe the data provided in the messages via voice recognition, and how the voice processing information is obtained from the web interaction and communications from the EMS and transcriber’s computer terminal connected in a configuration that involves the communication party routed via network 204 to EMS 206 where such a network is not limited to an Intranet, the Internet or any other communications network.

Because in Polcyn, two or more of the plurality of web interactions modes are used independently or concurrently to retrieve speech processing information directly from the internet as explained above since Polcyn discloses how web interaction modes can be used independently or concurrently according to the plurality of orders of communication sessions where for instance a first segment may comprise information communicated from a communicating party in a first communication session to EMS 206, while a second segment may

comprise information communicated in a second session to EMS 206 in both instances the messages comprising various different types of data and how the plurality of web interaction modes encountered in the plurality of segments of the plurality of messages in the communication interactions are able to work in agreement with a transcription interface application via voice recognition, and how the voice processing information is obtained from the web interaction and communications from the EMS and transcriber's computer terminal connected in a configuration that involves the communication party routed via network 204 to EMS 206 where such a network is not limited to an Intranet, the Internet or any other communications network, and thus, Applicant's argument is not persuasive.

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claims 1, 4, 6, 8, 10, 14, 35, 38, 45-49 are rejected under 35 U.S.C. 103(a) as being unpatentable over Polcyn (U.S. Patent: 6,865,258) in view of Cox et al. (U.S. Patent: 6,192,339) and further in view of Sibal et al. (U.S. Patent Application 2003/0182622), hereinafter referred to as Polcyn, Cox and Sibal.

With respect to **Claims 1, 35, 45**, Polcyn discloses:

A method, **non-transitory** machine-readable medium having instructions which when executed cause a machine to, and system (**Method, System and Computer Readable Medium, Polcyn,**

Col. 8, Line 62-Col. 9, Line 12, Col. 17, Lines 47-60, Col. 18, lines 1-13, and see also Cox, Col. 3, Lines 13-35) comprising:

receiving at a server computer system a client request from a client computer device via a network (**calling party transcription request and EMS 206 in communication via network 204, Col. 7, Lines 56- Col. 8, Line 28**);

interpreting the client request including identifying a selection of at least one of a plurality of web interaction modes (EMS 206 may comprise voice capture capability, voice record capability, voice play capability, voice recognition, DTMF recognition, Col. 10, Lines 10-19), each of the plurality of web interaction modes to perform interpretation of content being transmitted between the server computer system and the client

computer device, **wherein two or more of the plurality of web interaction modes are used independently or concurrently to retrieve speech processing information directly from the Internet** (the communicating party routed via network 204 to EMS 206 where EMS 206 comprises capability to receive image, fax, video, email; various forms of data may be communicated to the EMS 206 such as audio data, DTMF data, fax data, textual data, Col. 10, Lines 6-19, 37-61, and how for instance in Col. 11, Lines 17-60 of the Polcyn reference, it is disclosed how a message may comprise various different types of data such as a segment may comprise different types of data communicated to EMS 206 where the message may be thought of as a “container” in which various media are generated and organized into segments of different types. Furthermore, it is also disclosed in Polcyn how web interaction modes can be used independently or concurrently according to the plurality of orders of communication sessions where for instance a first segment may

comprise information communicated from a communicating party in a first communication session to EMS 206, while a second segment may comprise information communicated in a second session to EMS 206 in both instances the messages comprising various different types of data. At Col. 12, Lines 53-65 and Col. 10, Lines 6-36, of Polcyn for example, it is also disclosed how the plurality of web interaction modes encountered in the plurality of segments of the plurality of messages in the communication interactions are able to work in agreement with a transcription interface application that may automatically transcribe the data provided in the messages via voice recognition, and how the voice processing information is obtained from the web interaction and communications from the EMS and transcriber's computer terminal connected in a configuration that involves the communication party routed via network 204 to EMS 206 where such a network is not limited to an Intranet, the Internet or any other communications network); and

identifying a web interaction mode selected by the client computer device (the transcription interface application monitors the transcriber's activity and automatically adjusts the presentation of data to be transcribed according to such activity data type, Col. 12, Lines 18-35), and performing speech processing based on the selected web interaction mode and the retrieved speech recognition processing information (the transcription application utilizes voice recognition where the segment may be automatically transcribed and displayed in the appropriate field of data entry screen, Col. 14, Lines 14-32, 48-59; see also how the plurality of web interaction modes encountered in the plurality of segments of the plurality of messages in the communication interactions are able to work in agreement with

a transcription interface application via voice recognition, and how the voice processing information is obtained from the web interaction and communications from the EMS and transcriber's computer terminal connected in a configuration that involves the communication party routed via network 204 to EMS 206 where such a network is not limited to an Intranet, the Internet or any other communications network, Col. 12, Lines 53-65 and Col. 10, Lines 6-36, Col. 11, Lines 17-60), wherein performing speech processing includes determining an active display element that is to be focused (transcription application determines the transcriber's focus by determining the position of the cursor, Col. 17, Lines 21-39) and identifying the active display element with its associated identifier (the transcription application identifies the appropriate message segment corresponding to the transcriber's focus at block 410, and the transcription application may begin the presentation of the data of the appropriate message segment, Col. 17, Lines 21-39).

Polcyn, however, does not explicitly disclose, but Cox discloses **wherein the active display element includes an element upon which a speech input received from a user is focused, the speech input is received via the client computer device** (Depending on the voice input, corresponding machine commands derived from transition command mapping 216 are issued and the appropriate speech applications become focused and begin executing...a user may say: "switch to device control program" through microphone 402...", and also see how transition command mapping 216 may utilize different semantics for its generated statements, such as "focus on application XYZ" or "execute application XYZ". Also, transition command mapping 216 may display to a local user a list of available speech applications to choose from, Cox, Col. 5, Lines 59-67, Col. 6, Lines 40-64, Col. 5, Lines 4-18, client computer device in

Figure 4 and Figure 4 as a whole, Figure 3, elements 314, 316, Figure 2, central information object 200 and distributed computer systems 100 with remote usage tractability and servicing as server),

receiving an utterance from a user, via the client computer device, once the active display element is focused (receiving voice input where transition command mapping 216 may utilize different semantics for its generated statements, such as "focus on application XYZ" or "execute application XYZ", Cox Col. 5, Lines 4-18, Figure 4 as a whole and also Figure 3, elements 314, 316), **and, if the utterance matches the speech input, transmitting the identifier to the server computer system so that speech recognition is performed** (*"...support remote usage for both applications for the remote capability attribute..."*, *"...central information object 200..."* according to voice input matching of the focused application XYZ, Col. 6, lines 18-33, Col. 7, lines 21-48, Col. 5, Lines 4-18, client computer device in Figure 4 and Figure 4 as a whole, Figure 3, elements 314, 316, Figure 2, central information object 200 and distributed computer systems 100 with remote usage tractability and servicing as server),

performing speech recognition based on a relationship between the active display element and one or more speech elements (*"...central information object 200 maintains information from listeners 202 and speech applications such as device control program 104 and answering machine program 106...Speech applications may either modify or retrieve the information stored in central information object 200 through signaling interface 206...Central information object 200 may contain any of the following data, but not limited to, 1) currently focused speech application, 2) listening state of any speech recognition engine, 3) performance parameters and 4) graphical user interface support information. Multiple speech applications utilize these data*

to comprehend the running states of MASE 102...results in their seamless interactions with one another...”, “...device control program 104 and answering machine program 106...” ; “...transition command mapping 216 at this point likely contains “switch to device control program” and “switch to answering machine program” items. Depending on the voice input, corresponding machine commands derived from transition command mapping 216 are issued and the appropriate speech applications become focused and begin executing...a user may say: “switch to device control program” through microphone 402...”, “...multiple listeners 202 to be active simultaneously, but limit to a single instance of speech application per listeners 202. As a result, multiple distinct speech applications can coexist simultaneously...all applications have access to the state information of other applications and the environment...”, Col. 3, lines 60-67- Col. 4, lines 1-8, lines 54-67, Col. 5, lines 36-52, Col. 7, line 53-Col. 8, line 6, Col. 5, Lines 59-67, Col. 6, Lines 40-64, Figure 4 as a whole and also Figure 3, elements 314, 316, Figure 2, central information object 200 and distributed computer systems 100 with remote usage tractability and servicing).

Polcyn and Cox are analogous art because they are from a similar field of endeavor in facilitating improved web accesses applications via speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Polcyn with the technique for commanding visual and voice browsers in a common development platform and common environment taught by Cox in order to advantageously provide the user the desirability to dictation functionality in one product and device control functionality in another simultaneously in a seamless fashion, (Cox, Lines 51-55).

Polcyn in view of Cox, however does not explicitly disclose, but Sibal discloses **wherein performing speech recognition includes** retrieving a synchronization relationship between the one or more speech elements and the active display element to compose grammar of the one or more speech elements (**synchronizing field/partial field inputs between voice and visual browsers so that the user can fill out different fields of a single form using a combination of both voice and visual/tactile mode; synchronizing the voice browser by pointing the voice browser to a dialog on the VXML page that corresponds to that field; “granularity”; multi-modal platform 110 communicatively connected to and from computer device 102 and Web server 120 storing and/or generating markup content, Paragraphs [0031], [0032], [0040], [0045], [0055]-[0058], [0135], Figs. 1, 5 and 7), and dynamically correcting the composed grammar of the one or more speech elements using a real-time speech recognition based on the synchronization relationship (**field/partial field inputs allowing the user to type “New” and speak “York”; typing the city “New York” and speaking the zip code “10001” according to “granularity”, Paragraphs [0031], [0032], [0033], [0034], [0055]-[0058], [0135]**).**

Polcyn, Cox and Sibal are analogous art because they are from a similar field of endeavor in facilitating improved web accesses applications. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Polcyn in view of Cox with the technique for synchronizing visual and voice browsers to enable multi-modal browsing taught by Sibal in order to advantageously provide the user the usability to both browsers (visual and voice browsers) to interact with content simultaneously, (Sibal, Paragraphs [0004]-[0006]).

With respect to **Claims 4 and 38**, Sibal discloses:

wherein the focused active element comprises a hyperlink or a field in a form **(the user can fill out a single field using a combination of voice and visual/tactile input (e.g., entering a city name by typing “New” followed by speaking “York”))**, Paragraphs [0031], [0032], [0025], [0027], [0055]-[0058], [0135]).

With respect to **Claims 6 and 46**, Sibal discloses:

further including: extracting speech features from a user speech input, wherein the user speech input is contained in the client request **(synchronizing field/partial field inputs between voice and visual browsers so that the user can fill out different fields of a single form using a combination of both voice and visual/tactile mode**, Paragraphs [0031], [0032], [0040], [0020], [0022], [0024], [0025], [0027], [0055]-[0058], [0135], Figs. 1, 5 and 7).

With respect to **Claims 8 and 47**, Sibal discloses:

further including: receiving a session message at the server computer system to initialize a connection between the server computer system and the client computer device, wherein the session message includes an internet protocol (IP) address of the client computer device, a device type of the client computer device, a voice character of a user responsible for the user speech input, a language of the user input, and a default recognition accuracy requested by the client computer device **(multi-modal platform 110 communicatively connected to computer device 102 and Web server 120, client/server topology, web page 106 as portal page allowing the client to send request; computer device 102 requesting according to HTTP protocol; type of device; playing audio through speaker; multi-modal platform 110 configured to a “hit”**

of its own port as a signal to send information to visual browser, Paragraphs [0020]-[0022], [0024], [0025], [0027], [0028], [0031], 0032], [0038], [0055]-[0058], [0135]).

With respect to **Claims 10 and 48**, Sibal discloses:

further including: receiving a transmission message at the server computer system to exchange transmission parameters between the server computer system and the client computer device **(multi-modal platform 110 communicatively connected to computer device 102 and Web server 120, client/server topology, web page 106 as portal page allowing the client to send request, Paragraphs [0020], [0022], [0024], [0025], [0027])**.

Also, Polcyn disclose the communicating party routed via network 204 to EMS 206 where EMS 206 comprises capability to receive image, fax, video, email; various forms of data may be communicated to the EMS 206 such as audio data, DTMF data, fax data, textual data, (Col. 10, Lines 6-19, 37-61).

With respect to **Claims 14 and 49**, Sibal discloses:

further including: receiving an exit message at the server computer system to terminate a user session with the server computer system and the client computer device **(logger module, time stamping, Paragraph [0278])**.

Conclusion

8. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. See Form PTO-892.
9. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

10. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Edgar Guerra-Erazo whose telephone number is (571) 270-3708. The examiner can normally be reached on M-F 7:30a.m.-5:00p.m. EST. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, James Wozniak can be reached on (571) 272-7632. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Art Unit: 2626

/Edgar Guerra-Erazo/
Examiner, Art Unit 2626

/James S. Wozniak/
Supervisory Patent Examiner, Art Unit 2626